

# Speech Processing 2010/11

## 1st Test

March 24th 2011

Please identify this form with your name and student number in the reserved space at the bottom. The answers to multiple-choice questions will only be accepted if inserted in the appropriate place. Wrong answers will be penalized. The phonetic symbols should use the SAMPA alphabet (Lisbon accent).

1. Classify as True (T) or False (F)

- (a) Pitch is not perceptible if the first harmonic is not present.
- (b) The mel scale involves an arctan function.
- (c) The cepstral coefficients of higher order represent the filter contribution.
- (d) Low values of the zero crossing rate correspond to vowel segments.
- (e) The autocorrelation method of linear prediction analysis minimizes the error in a finite window.
- (f) Pre-emphasis is used to compensate the typical decaying of speech signals in unvoiced regions.
- (g) Linear prediction coefficients are not very sensitive to quantization.
- (h) The ratio of cross sectional areas of a lossless tube equivalent to the vocal tract is modeled in LSP coefficients.
- (i) PLP coefficients use a cubic root to perform intensity / loudness conversion.
- (j) A window of length  $N=40$  ( $F_s=8000$ ) may be used for computing the pitch value of a male voice segment, using the autocorrelation method.

2. Give examples of sounds for European Portuguese with the following properties (insert “-” in case they do not exist). Only one example is required for each category.

- (a) fricative unvoiced palatal
- (b) plosive labial unvoiced
- (c) nasal velar consonant
- (d) Nasal back vowel
- (e) non-nasal glide

3. Which of the 4 pictures in Figure 3 corresponds to

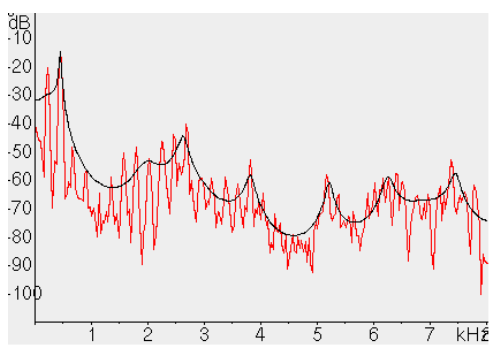
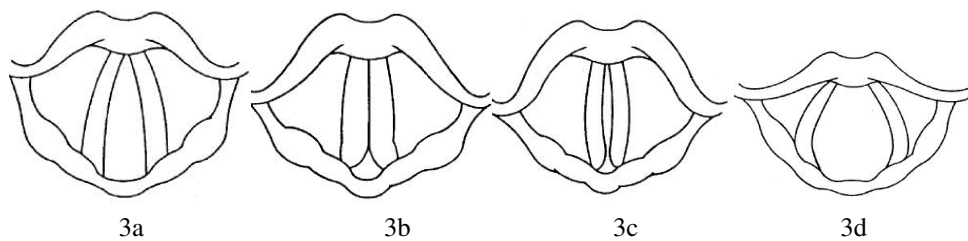
- (a) deep inhaling
- (b) phonation
- (c) whisper
- (d) normal breathing

4. Identify the 10-digit sequence by inspecting the spectrograms and waveforms in pages 3 and 4. There are no repeated digits. The recordings correspond to telephone speech.

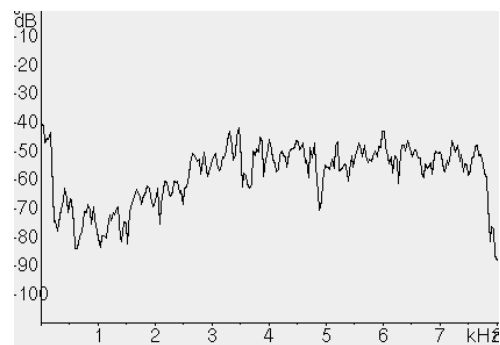
5. The spectrograms correspond to narrow band or wide band spectrograms?

6. If the second formant of a neutral vowel is 1800 Hz, what is the first formant and what is the corresponding vocal tract length?

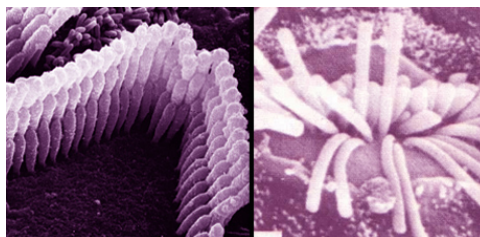
7. Write the expression for the real cepstrum  $c[n]$  of a signal  $x[n]$ , whose Fourier transform is given by  $X(e^{jw})$ .
8. Figure 8a presents the short-time Fourier transform (magnitude) of a vowel segment, including its LPC spectral envelope.
- Indicate approximate values for  $F_0$ ,  $F_1$  and  $F_2$  (Hz).
  - To which of the 3 vowels of the extremes of the vowel triangle does it correspond?
  - May it correspond to a window of 5 ms?
  - May it correspond to a male voice?
  - May the spectral envelope be obtained with an LPC analysis of order 8?
  - Does it correspond to telephone speech?
  - What class of sound (in terms of manner of articulation) may correspond to the short-time Fourier transform shown in Figure 8b?
9. Consider the following sentence:  
O Estado passa a cobrar uma percentagem dos montantes de receitas  
próprias angariados por professores do IST.
- Write the broad phonetic transcription.
  - What are the missing fricatives in this transcription?
10. Figures 10 show undamaged and damaged hair cells, caused by loud environmental noise, recreational activities, music and toys that exceed safe levels of sound. Which figure shows the undamaged hair cells? (This question is not for evaluation.)



8a

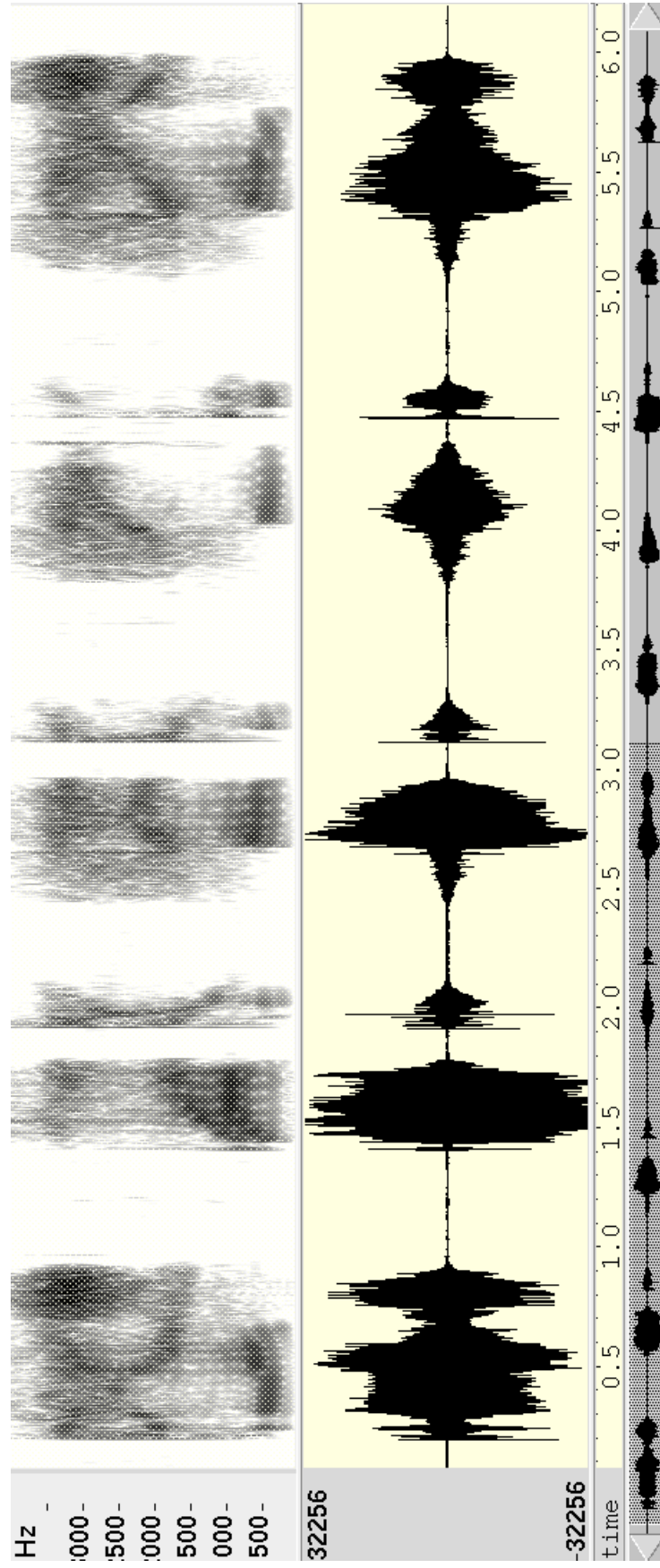


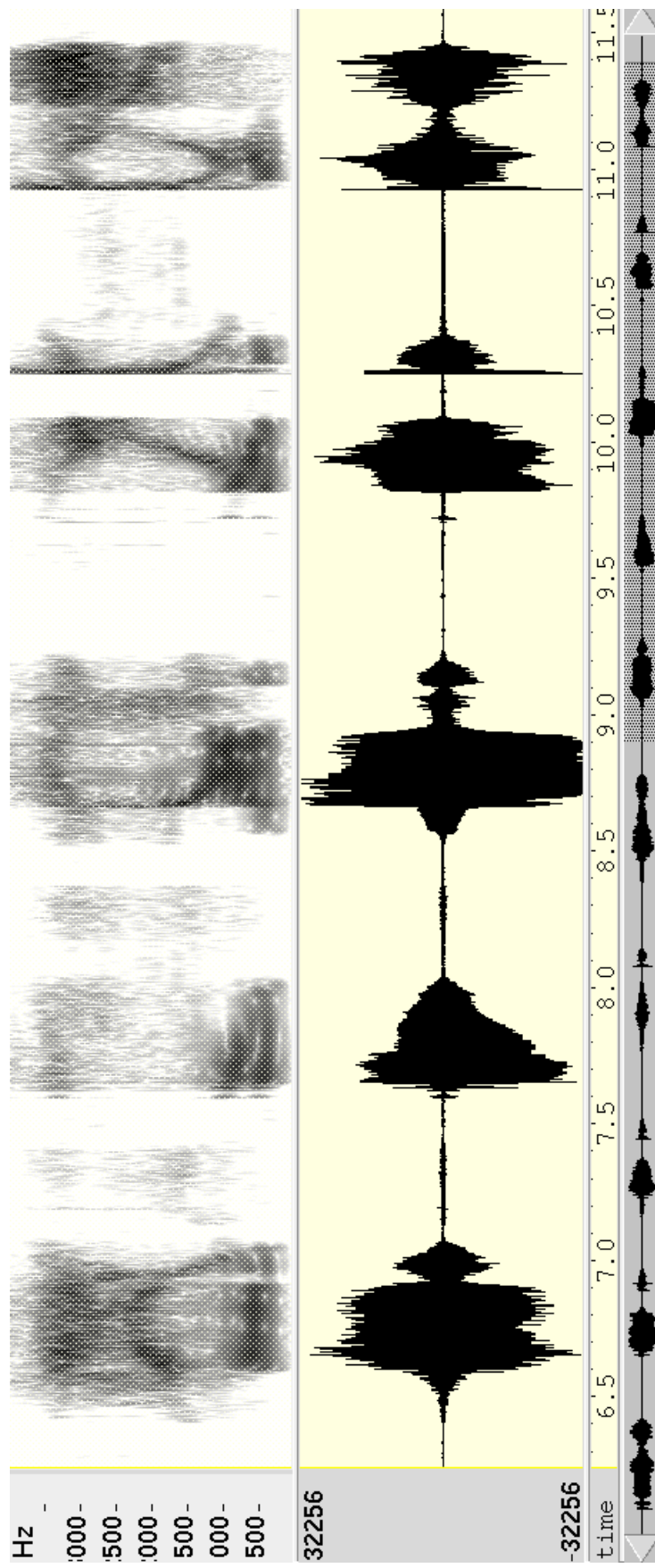
8b



10a

10b





**Test 1 - Answers**

Name:	
Number:	

1. (2.0 val.) Indicate T or F:

a	b	c	d	e	f	g	h	i	j

2 (2.0 val.)

a	
b	
c	
d	

3. (2.0 val.)

a	
b	
c	
d	

4. (3 val.)

--	--	--	--	--	--	--	--	--	--

5 to 7. (0.5/2.0/1.5)

5	
6	
7	
10	

8. (4.8 val.) Complete:

F0	F1	F2	Vow	Y/N	Y/N	Y/N	Y/N	Manner

9. (1.8/0.4 val.)

(a)	
(b)	